

УДК 519.876

Павло Струбицький, Ірина Струбицька

Тернопільський національний економічний університет, Україна

ПОБУДОВА ДИСКРЕТНИХ МОДЕЛЕЙ ЯК ПРОЦЕС МАШИННОГО НАВЧАННЯ

Pavlo Strubytsky, Iryna Strubyska

DISCRETE MODELS CONSTRUCTION AS MACHINE LEARNING PROCESS

Важливою властивістю будь-якої моделі, яку вона повинна набути в процесі навчання, є її здатність до узагальнення. Якщо модель навчилася і набула цю властивість, то вона буде видавати правильний результат при подачі на її вхід не тільки даних, на яких вона навчалася, але й даних, які не приймали участі в процесі навчання.

По суті, оптимізація параметрів при побудові дискретних динамічних моделей є машинним навчанням. Програма навчається на прикладі, а після закінчення фази навчання може узагальнювати. Тобто система не просто вивчає наведений приклад, а й вивчає закономірності в даних для навчання. При певному наборі параметрів модель повинна відтворювати закономірності, отримані під час навчання, для подальшого прогнозування.

У теорії нейронних мереж, щоб перевірити здатність моделі до узагальнення, всю навчальну вибірку поділяють на дві множини [1]:

- 1) навчальна;
- 2) тестова.

Навчальна вибірка використовується безпосередньо для навчання моделі. Тестова не використовується для навчання, а для перевірки властивості узагальнення моделі.

Тоді у процесі навчання моделі можна виділити дві похибки:

- 1) похибка навчання;
- 2) похибка узагальнення.

Похибка навчання — це помилка, яка допущена моделлю на навчальній вибірці. На кожній ітерації навчання для неперервної вхідної змінної вона розраховується як середньоквадратична похибка.

Похибка узагальнення — це помилка, яка отримала на тестових прикладах, тобто обчислюється по тих же формулах, але для тестової множини [2].

Якщо в процесі навчання була отримана достатньо мала похибка як на навчальній, так і на тестовій вибірці, то можна зробити висновок, що модель набула властивість узагальнення.

Розміри навчальної і тестової вибірки залежать від конкретної задачі. Навчальна множина повинна бути більша, ніж тестова, і містити достатньо даних для якісного навчання моделі. Розмір тестової множини визначається запасом даних навчальної вибірки. Якщо даних мало, то тестову вибірку беруть по менше — 5 % від загального об'єму вибірки. Якщо навчальна вибірка має достатній запас, то тестову множину можна брати 20–30 % [3].

Пропонуємо застосувати даний підхід при моделюванні дискретних динамічних систем, так як і при їх побудові модель спочатку навчається на вхідних даних. Вхідні дані моделі розіб'ємо на дві вибірки (навчальну і тестову) і обчислимо відповідні похибки на кожній ітерації оптимізаційного підходу.

Для наочності побудуємо макромодель трьохполюсника, вхідними значеннями якого є значення напруги, а вихідними — сили струму. Кількість дискретів рівна 58. Отже, навчальна вибірка буде складатись з 55 значень, а тестова, відповідно, - 3.

Для навчальної вибірки значення похибки зображено на рис. 1, а для тестової — на рис. 2.

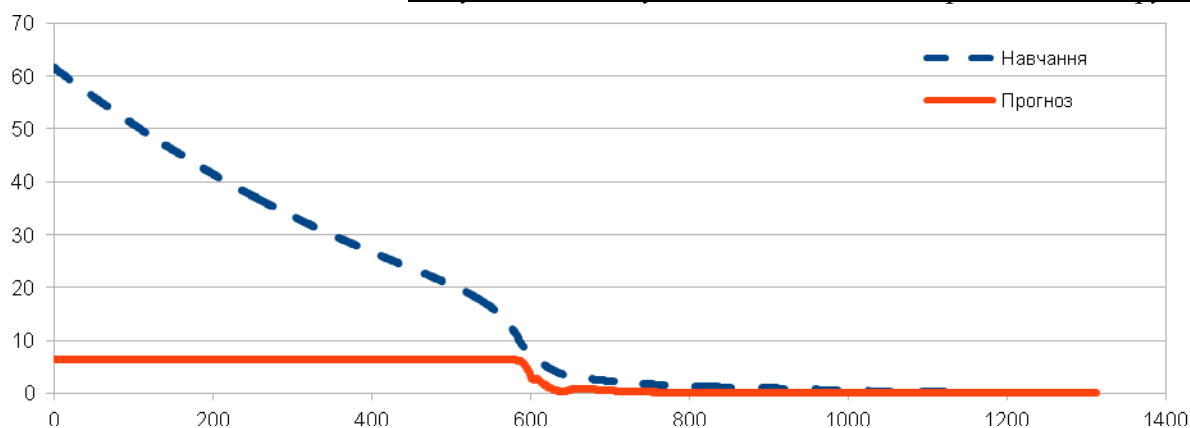


Рис. 1. Похибка на навчальній вибірці

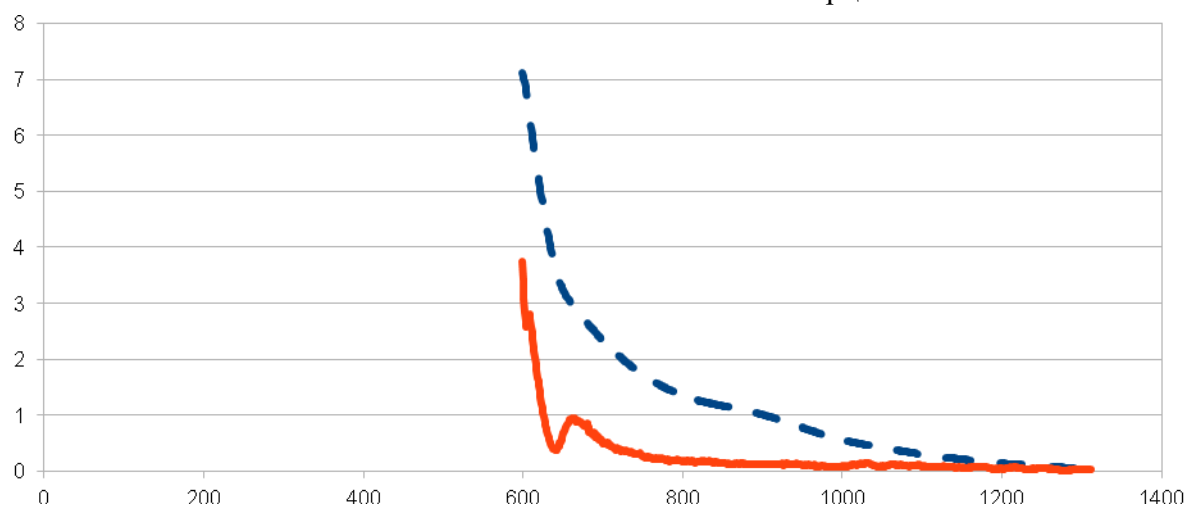


Рис. 2. Похибка на тестовій вибірці

Процес навчання моделі є задачею мінімізації функції вихідної похибки $E(t)$, де t — номер ітерації навчання. В процесі навчання похибка навчання монотонно зменшується. В той же час після певної ітерації може спостерігатись збільшення помилки узагальнення, що говорить про погіршення властивості узагальнення моделі. Тобто виникає, так званий, ефект перенавчання.

Виходячи з вище написано, можна зробити висновок, що інколи процес навчання потрібно зупиняти, це дасть змогу:

- 1) зменшити похибку узагальнення моделі;
- 2) скоротити кількість ітерацій і відповідно зменшити час при побудові моделі.

Література

1. Alpaydin E. Introduction to Machine Learning / Ethem Alpaydin. Cambridge, MA: The MIT Press, 2010. — 579 p.
2. Bishop C. Pattern Recognition and Machine Learning. Information Science and Statistics / Christopher M. Bishop. - New York: Springer, 2009. — 738 p.
3. Mitchell T. Machine Learning / Tom M. Mitchell. New York: McGraw-Hill, 1997. — 414 p.